

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ БІОРЕСУРСІВ І ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ

Кафедра комп'ютерних наук



РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ ІНТЕЛЕКТУАЛЬНИЙ АНАЛІЗ ДАНИХ НА ОСНОВІ МЕТОДІВ ШТУЧНОГО ІНТЕЛЕКТУ

Спеціальність – 122 «Комп'ютерні науки та інформаційні технології»

Факультет інформаційних технологій

Розробник: с.н.с., канд. фіз.-мат. наук, Лялецький О.В.

Київ – 2019

1. ОПИС НАВЧАЛЬНОЇ ДИСЦИПЛІНИ **«ІНТЕЛЕКТУАЛЬНИЙ АНАЛІЗ ДАНИХ НА ОСНОВІ** **МЕТОДІВ ШТУЧНОГО ІНТЕЛЕКТУ»**

Галузь знань, напрям підготовки, спеціальність, освітньо-кваліфікаційний рівень	
Галузь знань	12 “Інформатика та обчислювальна техніка”
Напрямок підготовки	122 “Комп’ютерні науки та інформаційні технології”
Освітній ступінь	бакалавр
Характеристика навчальної дисципліни	
Вид	Вибіркова
Загальна кількість годин	120
Кількість кредитів ECTS	4
Кількість змістових модулів	
Курсовий проєкт (робота) (якщо є в робочому навчальному плані)	
Форма контролю	Іспит
Показники навчальної дисципліни для денної та заочної форм навчання	
	денна форма навчання
Рік підготовки	4
Семестр	8
Лекційні заняття	12 год.
Практичні, семінарські заняття	
Лабораторні заняття	24 год.
Самостійна робота	84 год.
Кількість тижневих годин для денної форми навчання: аудиторних самостійної роботи студента –	3 год.

2. МЕТА ТА ЗАВДАННЯ НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Навчальна дисципліна «Інтелектуальний аналіз даних на основі методів штучного інтелекту» є складовою частиною циклу базових дисциплін, які забезпечують підготовку бакалаврів за фахом «Комп'ютерні науки».

Метою і завданням навчальної дисципліни «Інтелектуальний аналіз даних на основі методів штучного інтелекту» є виклад основних понять, моделей і методів штучного інтелекту, які можуть використовуватися при вирішенні практичних завдань з інтелектуального аналізу даних, пов'язаних з класифікацією та прогнозуванням, розпізнаванням образів, прийняттям рішень в різних областях людської діяльності, в тому числі, сільськогосподарській, за допомогою комп'ютерної техніки.

Викладання дисципліни ґрунтується на знаннях по таких напрямках як:

- дискретна математика,
- структури даних;
- організація баз даних та знань;
- системи і технології управління базами даних;
- інформаційно-управляючі системи та технології в АПК.

Матеріал, який викладається у цій дисципліні, може бути використаним бакалаврами при написанні дипломних робіт.

3. ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Змістовий модуль 1. Основи інтелектуального аналізу даних.

Тема лекційного заняття 1. Розвиток і призначення інтелектуального аналізу даних. Основні поняття. Концепції, архітектура і класифікація систем.

Основні поняття та визначення інтелектуального аналізу даних. Концепції, архітектура і класифікація систем інтелектуального аналізу даних, області їх застосування. Основні задачі інтелектуального аналізу даних. Інтелектуальний аналіз даних і бази даних. Інтелектуальний аналіз даних і штучний інтелект. Можливі напрямки використання інтелектуального аналізу даних. [1,2,3]

Тема лекційного заняття 2. Процеси, активності, задачі і методи інтелектуального аналізу даних.

Характеристики процесів і активностей інтелектуального аналізу даних. Характеристики методів інтелектуального аналізу даних. Методи, засновані на збережених даних. Методи, засновані на дистиляції даних. Логічні методи. Статистичні методи. Інші методи. [2,3,5]

Тема лекційного заняття 3. Асоціативні правила і дерева рішень. Методи пошуку асоціативних правил та технологія застосування дерев рішень. [2,4]

Змістовий модуль 2. М'які обчислення в інтелектуальному аналізі даних

Тема лекційного заняття 4. Нечіткі обчислення та міркування в інтелектуальному аналізі даних.

Нечіткі множини, операції над нечіткими множинами. Поняття нечіткої і лінгвістичної змінної. Нечіткі висловлювання та нечіткі предикати. Правило нечіткого логічного виводу. Композиційне правило нечіткого виводу Заде. Використання нечіткого виводу в інтелектуальному аналізі даних. [1,5,6]

Тема лекційного заняття 5. Нейрокомп'ютерні обчислення.

Поняття нейрона, нейромережі, нейрокомп'ютера. Класифікація та види моделей нейромереж. Загальний опис розроблення, навчання та використання нейромереж. Використання нечітких гібридних нейронних мереж для побудови певних баз знань. [1,2,7]

Тема лекційного заняття 6. Еволюційне програмування, теоретичні основи генетичних методів та алгоритмів.

Теоретичні основи сучасного еволюційного програмування. Аналіз методів еволюційного програмування. Генетичні алгоритми, їх моделі. Переваги, недоліки і шляхи підвищення ефективності генетичних алгоритмів. [1,2,3].

4. Структура навчальної дисципліни

«Інтелектуальний аналіз даних на основі методів штучного інтелекту»

Назви змістових модулів і тем	Кількість годин												
	денна форма							Заочна форма					
	тижн і	усьог о	у тому числі					усього	у тому числі				
			л	п	лаб	інд	с.р.		л	п	лаб	інд	с.р.
1	2	3	4	5	6	7	8	9	10	11	12	13	14
Змістовий модуль 1. Основи інтелектуального аналізу даних. Класифікація систем та базові методи.													
Тема 1. Розвиток і призначення інтелектуального аналізу інформації. Концепції, архітектура і класифікація систем.	1	10	2		4		4						
Тема 2. Процеси, активності і методи інтелектуального аналізу даних.	1	22	2		4		16						

Тема 3. Асоціативні правила і транзакції. Методи пошуку асоціативних правил.	1	22	1		4		16						
Разом за змістовим модулем 1	3	54	6		12		36						
Змістовий модуль 2. М'які обчислення в інтелектуальному аналізі даних.													
Тема 4. Нечіткі обчислення та міркування в інтелектуальному аналізі даних.	1	22	2		4		16						
Тема 5. Нейрокомп'ютерні обчислення.	1	22	2		4		16						
Тема 5. Еволюційне програмування, теоретичні основи генетичних методів та алгоритмів.	1	22	2		6		16						
Разом за змістовим модулем 2	3	66	7		18		48						
Усього	6	120	12		24		84						

5. ТЕМИ СЕМІНАРСЬКИХ ЗАНЯТЬ

Не передбачено програмою

6. ТЕМИ ПРАКТИЧНИХ ЗАНЯТЬ

Не передбачено програмою

7. ТЕМИ ЛАБОРАТОРНИХ ЗАНЯТЬ

№ з/п	Назва теми	Кількість годин
1	Теоретичні аспекти побудови систем інтелектуального аналізу даних. Основні підходи до технологій їх проектування та розробки. Етапи проектування, області застосування.	2 години
2	Розбір на конкретних прикладах методів, що засновані на збереженні та дистиляції даних.	2 години

3	Вирішення задач на застосування різних методів пошуку асоціативних правил.	4 години
4	Побудова нечітких множин, лінгвістичних змінних та простих правил продукцій. Рішення задач, пов'язаних з операціями над нечіткими даними та нечітким виводом.	6 години
5	Розбір блок-схеми процесу розроблення штучної нейромережі. Алгоритми навчання. Тестування і реалізація.	6 години
6	Розбір узагальненої схеми роботи генетичних методів. Оператор відбору. Оператор схрещування. Оператор мутації. Формування нового покоління. Критерії зупинення.	4 години

7. САМОСТІЙНА РОБОТА

Для кращого засвоєння матеріалу студентам пропонується самостійно опанувати наведеними нижче темами і, як результат освоєння відповідного матеріалу, підготувати реферати на ці теми.

1. Вивчення класифікації систем інтелектуального аналізу даних. Основні етапи проектування систем інтелектуального аналізу даних. Технології проектування та розробки систем інтелектуального аналізу даних. Знайомство з можливими напрямками використання інтелектуального аналізу даних. [1,2,3] (8 год.).

2. Знайомство з базовим підходами до інтелектуального аналізу даних, такими, як кластеризація, асоціація, послідовність, прогнозування, візуалізація, підведення підсумків. Знайомство з статистичними методами (аналіз часових рядів, кореляційно-регресійний аналіз дискримінантний аналіз, кластерний аналіз). [2,3,5] (16 год.).

3. Вивчення алгоритмів пошуку та обчислення асоціативних правил: алгоритму Apriori, алгоритму без генерування кандидатів, послідовних алгоритмів. Побудова та вибір дерев рішень, їх застосування. [2,4] (14 год.)

4. Вивчення основи теорії лінгвістичних змінних, розглянути приклади. Вивчення поняття нечіткого висловлювання та нечіткого предиката. Розв'язок задач з використанням операцій нечіткої логіки [3], [5]. Вивчення основних етапів нечіткого виводу: формування бази правил систем нечіткого виводу, фазифікація вхідних змінних, агрегування підумов, активізація підзаключень, акумулювання заключень. [1,5,6] (20 год.).

5. Вивчення таких типів нейронних мереж, як персептрон Розенблатта та багатошаровий персептрон, нейронні мережі Джордана, Елмана, Хеммінга, Хопфілда та Кохонена, когнітрон та неокогнітрон, хаотична нейронна мережа та осциляторна нейронна мережа, мережі узагальненої регресії та ймовірнісні мережі, сіамська нейронна мережа та мережа адаптивного резонансу. [1,2,7] (16 год.).

6. Вивчення моделей генетичних алгоритмів, мурашиний алгоритм. Вивчення методів побудови та реалізації генетичних алгоритмів. Використання генетичних алгоритмів для рішення задач класифікації. [1,2,3] (16 год.).

8. МЕТОДИ НАВЧАННЯ

М1. Лекція (проблемна, інтерактивна). М2. Лабораторна робота. М3. Проблемне навчання. М4. Проектне навчання (індивідуальне). М8. Дослідницький метод.

9. ФОРМИ КОНТРОЛЮ

МК1. Тестування. МК4. Методи усного контролю. МК5. Екзамен. МК7. Звіт

10. РОЗПОДІЛ БАЛІВ, ЯКІ ОТРИМУЮТЬ СТУДЕНТИ

Поточний контроль				Рейтинг з навчальної роботи $R_{нр}$	Рейтинг з додаткової роботи $R_{др}$	Рейтинг штрафний $R_{штр}$	Підсумкова атестація (екзамен чи залік)	Загальна кількість балів
Змістовий модуль 1	Змістовий модуль 2	Змістовий модуль 3	Змістовий модуль 4					
0-100	0-100	0-100	0-100	0-70	0-20	0-5	0-30	0-100

Примітки. 1. Відповідно до «Положення про кредитно-модульну систему навчання в НУБіП України», затвердженого ректором університету 03.04.2009 р., рейтинг студента з навчальної роботи $R_{нр}$ стосовно вивчення певної дисципліни визначається за формулою

$$R_{нр} = \frac{0,7 \cdot (R^{(1)}_{зм} \cdot K^{(1)}_{зм} + \dots + R^{(n)}_{зм} \cdot K^{(n)}_{зм})}{K_{дис}} + R_{др} - R_{штр},$$

де $R^{(1)}_{зм}, \dots, R^{(n)}_{зм}$ – рейтингові оцінки змістових модулів за 100-бальною шкалою;

n – кількість змістових модулів;

$K^{(1)}_{зм}, \dots, K^{(n)}_{зм}$ – кількість кредитів ECTS, передбачених робочим навчальним планом для відповідного змістового модуля;

$K_{дис} = K^{(1)}_{зм} + \dots + K^{(n)}_{зм}$ – кількість кредитів ECTS, передбачених робочим навчальним планом для дисципліни у поточному семестрі;

$R_{др}$ – рейтинг з додаткової роботи;

$R_{штр}$ – рейтинг штрафний.

Наведену формулу можна спростити, якщо прийняти $K^{(1)}_{зм} = \dots = K^{(n)}_{зм}$. Тоді вона буде мати вигляд

$$R_{нр} = \frac{0,7 \cdot (R^{(1)}_{зм} + \dots + R^{(n)}_{зм})}{n} + R_{др} - R_{штр}.$$

Рейтинг з додаткової роботи R_{др} додається до **R_{нр}** і не може перевищувати 20 балів. Він визначається лектором і надається студентам рішенням кафедри за виконання робіт, які не передбачені навчальним планом, але сприяють підвищенню рівня знань студентів з дисципліни.

Рейтинг штрафний R_{штр} не перевищує 5 балів і віднімається від **R_{нр}**. Він визначається лектором і вводиться рішенням кафедри для студентів, які матеріал змістового модуля засвоїли невчасно, не дотримувалися графіка роботи, пропускали заняття тощо.

2. Згідно із зазначеним Положенням **підготовка і захист курсового проекту (роботи)** оцінюється за 100 бальною шкалою і далі переводиться в оцінки за національною шкалою та шкалою ECTS.

Оцінювання студента відбувається згідно положенням «Про екзамени та заліки у НУБіП України» від 27.02.2019р. протокол №7

Оцінка національна	Рейтинг здобувача вищої освіти, бали
Відмінно	90-100
Добре	74-89
Задовільно	60-73
Незадовільно	0-59

11. МЕТОДИЧНЕ ЗАБЕЗПЕЧЕННЯ

1. Голуб Б. Л., Ящук Д. Ю. Організація сховищ даних: Навчальний посібник до вивчення дисципліни «Організація сховищ даних». – Київ, 2018. – 150 с.

12. РЕКОМЕНДОВАНА ЛІТЕРАТУРА

Базова

1. Ситник В.Ф., Краснюк М.Т. Інтелектуальний аналіз даних (дейта-майнінг). Київ: КНЕУ, 2007. – 376 с.
2. А. О. Олійник, С. О. Субботін, О. О. Олійник. Інтелектуальний аналіз даних. – Запоріжжя: ЗНТУ, 2012. – 278 с.
3. А. А. Шумейко, С. Л. Сотник. Интеллектуальный анализ данных (Введение в Data Mining): учеб. пособ. – Днепропетровск, 2012. – 212 с.

Додаткова

4. В. В. Ковальчук. Інтелектуальний аналіз даних (Конспект лекцій). – Одеса: ОДЕКУ, 2015. – 206 с.
5. Д.В. Ланде, І.Ю. Субач, Ю.Є. Бояринова. Основи теорії і практики аналізу інтелектуального аналізу даних у сфері кібербезпеки: навчальний посібник. – К.: ІСЗЗІ КПІ ім. Ігоря Сікорського», 2018. – 297 с.
6. Г.Э. Яхьяева Основы теории нечетких множеств: Курс Интернет-университета информационных технологий. 2016. – 187 с.
7. Саймон Хайкин. Нейронные сети: полный курс, 2-е издание. К: Диалектика-Вильямс 2005. – 1104 с.

13. ІНФОРМАЦІЙНІ РЕСУРСИ

1. <https://dl.sumdu.edu.ua/textbooks/85292/354092/index.html>